CHAPTER 6

The Measure of All Things: Quantum Mechanics and the Soul

Hans Halvorson

INTRODUCTION

The twentieth century saw several significant developments in our understanding of the physical world. One of the most significant of these developments was the replacement of the classical physics of Newton, Maxwell, and Einstein with the quantum physics of Planck, Bohr, and Heisenberg.

Usually our understanding of the universe grows at an agonizingly slow pace. For example, a group of scientists might spend years figuring out the next digit in the decimal expansion of some seemingly insignificant numerical parameter. Of course, every now and then, there is a discovery that finds its way into the popular consciousness. For example, scientists might discover a new object (e.g. a new star) or even a new type of object (e.g. a new species). But it is only on the rarest of occasions that an actual scientific revolution occurs, when an old theory (and its accompanying world picture) is dispensed with in favor of a new theory (with a new understanding of ourselves and our place in the universe). The introduction of quantum mechanics may be the greatest scientific revolution to date in human history: the replacement of classical physics by quantum physics requires a thoroughgoing modification of our world view; or as philosophers might say, it requires a modification of our fundamental metaphysics.

That much is clear. But there is little consensus about how to build a new world view around quantum mechanics. For example, some claim that quantum mechanics proves that the universe is indeterministic, and the future is open. Others claim more radically that quantum mechanics shows that there is a multitude of parallel universes, and that each time a measurement is made, our universe branches again. Still others claim that quantum mechanics proves that there is no objective world outside of our perceptions.

The main goal of this chapter is to put forward an alternative view of the metaphysical lessons of quantum mechanics. But let me begin by staking out my methodology: I do not believe that it is feasible to approach quantum mechanics from a standpoint of "metaphysical neutrality," and expect it to tell us the nature of the universe. Rather, we always approach scientific theories in light of our background beliefs; we can then ask if this theory is consistent with these beliefs, and whether or not it suggests modifications of these beliefs.[1] For example, these background beliefs might include the belief that there is an external world, or the belief that the universe did not come into existence (along with all of our memories) one second ago, or the belief that there are conscious persons besides myself.

One of the more controversial background beliefs that I bring to this investigation is the Soul Hypothesis — namely the belief that human beings are more than just their bodies, but are also living souls. I will argue that quantum mechanics says nothing to suggest that we must abandon the Soul Hypothesis. Indeed, I will show that the Soul Hypothesis allows us to reject some of the more wild and implausible metaphysical speculations based on quantum mechanics.

The remainder of this chapter will proceed as follows. In the second section, I give an informal sketch of quantum mechanics; in particular, I isolate four central features of the theory that give rise to various paradoxes. In the third section, I discuss a much more serious paradox, the so-called "measurement problem" of quantum mechanics. The measurement problem supposedly shows that an observer (like you or me) could not ascertain facts about the physical world by making observations, and so (among many other things) could not actually test quantum mechanics. In the fourth section, I briefly pause to discuss some popular resolutions of the measurement problem before returning, in the fifth section, to discuss the bearing of the Soul Hypothesis on the measurement problem.

BASIC ASSUMPTIONS OF QUANTUM MECHANICS

“Classical physics” is a catch-all phrase for a number of different theories developed roughly between the time of Galileo Galilei (1564–1642) and James Clerk Maxwell (1831–1879). Radically abstracting from the rich detail of these theories, they are all based on two main assumptions: first, the state of each object in the world can be completely specified by assigning values to all of that object’s quantitative properties (such as its position, its velocity, its mass, etc.). Second, there are laws of nature such that the state of each object at any future time is completely determined by the state of all objects at any previous time.

The classical physicists also successfully pursued a strategy of reductionism by finding a small number of “basic quantities” from which the values of all other quantities could (in principle) be determined. Famously, these basic quantities include things such as position and velocity, but exclude many quantities that figure centrally in our everyday lives, such as color and temperature.

How did these physicists know that they could not reduce the collection of basic quantities even further? For example, how did they know that velocity could not be reduced to position? They knew that velocity could not be reduced to position because these two quantities satisfy a *mix and match* principle. For example, the position and the velocity of a baseball can be mixed and matched in the sense that, in principle, the position of the baseball (e.g. over home plate) can be matched with any velocity of the baseball (e.g. traveling at 60 miles per hour). In contrast, the color of the baseball cannot be mixed and matched with the position and velocities of its constituent atoms; indeed, the color is completely determined by, or reducible to, the position and velocity of the constituent atoms.

During the late nineteenth century, physicists found ways to put classical mechanics to work even in cases where they lacked precise knowledge of the states of objects. In particular, given partial knowledge of the states of objects, the (deterministic) dynamical laws of the theory can be applied to yield partial knowledge about the future states of objects. Let’s consider a highly simplified example: suppose that there is a machine that releases a certain sort of classical particles, but that we do not have full control over the outgoing direction of these particles. Suppose then that the machine is confined to a box that has an optical screen at one end, and that between the machine and the screen there is another blocking screen that has two doors (see Figure 6.1). In each case when a particle is emitted, it will go through either the left door or

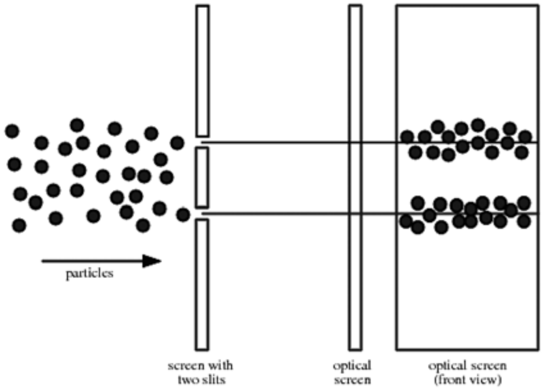

FIGURE 6.1 Particles

the right door; however, in any given case, we do not know which door will be traversed. Nonetheless, at the end of several runs of the experiment, it is overwhelmingly likely that there will be one "lump" on the optical screen behind the left door, and an equal sized lump behind the right door. In other words, after the particles pass through the doors, they follow the trajectories predicted by classical physics and so continue in a straight line to the optical screen. Extending the use of the word "state," we can say that this apparatus describes a state that is a probabilistic mixture of a state in which the particle goes through the left door and a state in which the particle goes through the right door.

Classical physics made another important advance when its domain was expanded from particles — i.e. well-localized discrete objects — to waves (as occur in media such as water and air) and fields (such as the electromagnetic field). One of the novel physical features of waves and fields is their ability to be superposed on top of each other. For example, if a certain wave machine at Waterworld produces 2-foot waves, and a certain other wave machine produces 3-foot waves, then if we set both machines in sync we would get 5-foot waves. In contrast, if we set the machines out of sync, then the peaks and troughs will interfere with each other so that we only get 1-foot waves. This special feature of waves (and fields) is called "superposability"; the wave that results from combining two other waves is called the "superposition" of those waves.

Of course, waves superpose in more than just one dimension. For example, suppose that we set up a source of monochromatic light (e.g. a laser) on one side of a box and an optical detector screen on the other side of the box. Suppose, moreover, that we place a barrier with two

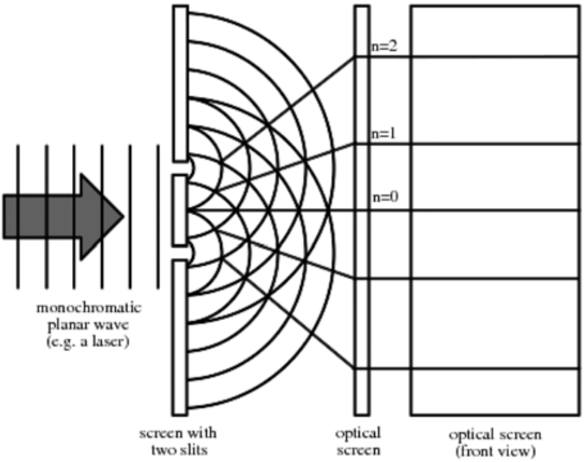

FIGURE 6.2 Waves

open doors in the middle of the box. Then as light waves come out of the individual doors, they will superpose with each other, reinforcing each other at some points and canceling each other out at other points, to form a characteristic diffraction pattern on the optical screen (see Figure 6.2).

SUPERPOSITION

Classical physics proved itself flexible enough to accommodate uncertainty (probability), and also to accommodate physical systems (such as waves) that are not composed of discrete particles. However, at the end of the nineteenth century, several new experiments provided data that could not be explained by classical physics. A striking example of these experiments is the famous two-slit experiment for electrons. In order to show that this experiment cannot be understood within classical physics, we must briefly recall the state of physical knowledge and of experimental technology at the end of the nineteenth century.

At the end of the nineteenth century, physicists had experimental data indicating that atoms exist, and moreover that atoms themselves are complex physical objects consisting of a dense nucleus with positive electric charge, and a less dense outer region with negative electric charge. But what is the negatively charged region made of? Is it made of small particles (viz. electrons) with empty space between them, or is it simply a continuous, infinitely divisible field of negative charge? Fortunately, experimentalists had devised methods of dislodging pieces of this negative charge, and so made it possible to perform experiments to test whether negative charge is carried by waves or particles.

Consider then an experiment in which there is a source of negative electric charge that is directed towards a detector screen, but that between the source of electrons and the detector, there is a barrier with two doors. The first experimental finding suggests that negative charge is carried by particles: if one of the two doors is open, and the other is closed, then the detector flashes only in the small region directly behind the open door. If electric charge were carried by a continuous, wavelike medium, then we would expect the charge to spread out after it passes through the door, and then to leave a broad, diffuse mark on the detector.

Now let's apply a second test: turn the source on for an extended period of time, and attach the two doors to a coin flipping machine (which opens the right door when the coin comes up tails, and opens the left door when the coin comes up heads). If the experiment is run several times, then the resulting pattern on the screen is exactly what we would expect for particles (as in Figure 6.1): there are two lumps of equal size, one behind each door.

Now let's apply a third test: turn the source on for an extended period of time, and open both doors. If electrons were localized particles, then they must pass through one of the two doors. Thus, if the source is set up so as not to bias one of the two doors, then over many runs of this experiment, a pattern would build up on the detector screen — one lump behind the left door, and another equal-sized lump behind the right door. But, in fact, that is *not* what happens in this experiment. Rather, at the end of the experiment, the detector screen displays the diffraction pattern that we saw in the two-slit experiment for waves (see Figure 6.2).

In the early days of quantum mechanics, a thought experiment was devised to try to settle the question of whether electrons are particles or waves: put a detector over each door and see if one, both, or neither detector goes off. Only very recently have technological advances made it possible to perform this experiment, and the result is surprising: in any particular run of the experiment when the detectors are turned on, exactly one detector goes off (confirming that electrons are localized particles). But when these detectors are turned on, the interference pattern on the optical screen disappears, and instead we get the two-lump pattern on the optical screen. It is as if the electron behaves like a particle in the presence of the detectors, but like a wave when there are no detectors.

So is negative electric charge carried by particles or waves? The pioneers of quantum mechanics refused to answer this question; instead, they constructed a hybrid theory that draws on features both of the classical theory of particles and of the classical theory of waves. In particular, they invented a new concept called the "quantum superposition of two states," and they claimed that when both doors are open, then the electron is in a quantum superposition of passing through the left and the right doors. If |*left*> is the state of the electron passing through the left door, and |*right*> is the state of the electron passing through the right door, then the quantum superposition state is usually written |*left*>+|*right*>.

In some ways, a quantum superposition behaves like a classical superposition of waves; e.g. when both doors are open but no detectors are turned on, then the quantum superposition also results in a diffraction pattern on the optical screen. But in some ways it does not; e.g. when both doors are open and both detectors are turned on, a classical superposition of waves would always set off both detectors at the doors, but a quantum superposition of waves will only set off one of the detectors at a time.

The key feature of the superposition state |*left*>+|*right*> is that it *cannot* be thought of merely as a state of our ignorance of which door the electron will pass through. That is, it cannot be thought of as a probabilistic mixture of the two states |*left*> and |*right*>. If it were merely a description of our ignorance, then there would be no diffraction pattern on the optical screen! And yet, the state |*left*>+|*right*> does predict what we would see if we were to look at which door the electron is passing through: we would see it go through the left door half of the time, and through the right door the other half of the time. In other words, when an electron is in state |*left*>+|*right*>, then it does not have any determinate position whatsoever, i.e. it is neither in the state |*left*> nor in state |*right*>. And yet, if we measure the position of the electron, e.g. by placing detectors over the doors, then there is a 50 per cent chance that the electron will change into state |*left*>, and a 50 per cent chance that it will change into state |*right*>. Since quantum states are often called "wave functions," this remarkable change of state has been given the infamous name, "collapse of the wave function."

Before we proceed further, it is *crucially important* to undercut a possible misunderstanding — a misunderstanding into which many professional physicists and philosophers have fallen. What are we to say about the condition of the electron before the wave function is collapsed, e.g. before we look at which door the electron is passing through?

Should we say that before a measurement is made, there is no reality, that the facts about physical reality are brought into existence by the act of measurement? No: such an idea is based on a complete misunderstanding of the formalism of quantum mechanics.

To clear up this misunderstanding, we need to point out that the states of quantum mechanics are *not* like the natural numbers, i.e. the numbers 1, 2, 3 . . . The natural numbers can be divided into two classes: the composite numbers (a number that can be divided by at least one number besides 1 and itself), and the prime numbers (those that are not composite). There is then a clear sense in which some numbers are simple and others are complex; the complex numbers can always be decomposed into simples, but the simples cannot be further decomposed.

But there is no similar distinction in quantum mechanics between states that are composite (superpositions), and states which are simple (not superpositions of other states). On the contrary, quantum states are like angles on a disc, or like points on the face of an analog clock, and the superposition operation "+" on quantum states is like taking the average (i.e. the midpoint going clockwise from the first angle) between the two points on the clock. For example, the states |*left*> and |*right*> are themselves superposition of other states, namely the state

|*moving*> = |*left*>+|*right*>,

in which the electron is moving, and the state

|*stationary*> = |*left*>–|*right*>,

in which the electron is stationary. There is no sense whatsoever in which some quantum states are *not* superpositions. As a result, there are no "safe" quantum states in which all "elements of reality" are fully determinate: in every quantum state, some quantities fail to have a determinate value.

But if every quantum state is a superposition (of some other states) then don't we have a serious reality crisis? Didn't we say that when an electron is in a superposition of states, then it fails to have the features specified by those states? Doesn't this mean that at any given time, some features of the electron will remain in a shadow land between existence and non-existence? Yes, if quantum mechanics is true, then an electron can *never*, in any state, have determinate values for all of its quantities. At any given time, an electron will either have no position, or no velocity, or no value for some other quantity.

Perhaps you can accept with equanimity the claim that electrons are in superposition states; after all, we cannot see them. But unfortunately for our grip on reality, it is not just subatomic particles that are in superpositions. As we will see in the next section, superpositions percolate upward in the sense that anything composed out of subatomic particles also has superposition states, indeed is always in a superposition of states. But rocks, trees, and human bodies are composed out of subatomic particles; and so we are always in a superposition of states!

ENTANGLEMENT

What happens if there are *two* electrons, both of which are in a superposition of states? Suppose, for example, that two separate two-slit experiments are performed in two different laboratories. If both electrons are in the state |*left*>+|*right*>, then what is the state of the composite of both electrons?

We can make some progress on this question by asking what we should expect to see if we were to measure the position of both electrons simultaneously. To keep track, let's give the two electrons names: Anke and Bert. If we simultaneously measure the position of both Anke and Bert, then since Anke is in the state |*left*>+|*right*>, there is a 50 per cent chance that she will go through the left door, and a 50 per cent chance that she will go through the right door. Similarly, there is a 50 per cent chance that Bert will go through his left door, and a 50 per cent chance that he will go through his right door. Now, supposing that these two experiments are independent from each other, the outcomes of the two measurements should satisfy the mix-and-match principle; that is, Anke's going through the left door is compatible with Bert's going through either the left or right door, and vice versa. Thus, we should expect to see each possible combination — left-left, left-right, right-right, right-left — 25 per cent of the times we do the experiment. And, indeed, that is the result that is observed when such experiments are performed.

Let us write $|\textit{left}>_{A}|\textit{left}>_{B}$ for the state in which both Anke and Bert go through their respective left doors. Then we began the discussion of this section by postulating that Anke and Bert are in the state:

$$(|\textit{left}>_{A}+|\textit{right}>_{A})(|\textit{left}>_{B}+|\textit{right}>_{B}),$$

in which both Anke and Bert are in the superposition |*left*>+|*right*>. Here we simply set the states side by side (with no space in between), to indicate that the state on the left belongs to Anke and the state on the right belongs to Bert. The evidence of joint measurements indicates that this state is in fact equal to a superposition of four terms,

$$|\mathit{left}>_A|\mathit{left}>_B+|\mathit{left}>_A|\mathit{right}>_B+|\mathit{right}>_A|\mathit{left}>_B+|\mathit{right}>_A|\mathit{right}>_{B.}$$

Notice that this superposition of four terms is what we would expect if we could distribute the superposition operation "+" over the composition of two systems. And indeed, quantum mechanics accepts the validity of distribution, for example:

$$|\mathit{moving}>_A(|\mathit{left}>_B+|\mathit{right}>_B) = |\mathit{moving}>_A|\mathit{left}>_B+|\mathit{moving}>_A|\mathit{right}>_{B.}$$

Again, this equation is completely plausible when you think about the results of measuring the velocity of Anke and the position of Bert. In particular, if Anke is definitely moving, and Bert is in a superposition of $|\mathit{left}>$ and $|\mathit{right}>$, then a joint velocity-position measurement will yield either "moving and left" or "moving and right."

Now, the composite of two electrons is still extremely small, and so certainly still within the domain of validity of quantum mechanics. In particular, the states of a pair of electrons should, theoretically speaking, be superposable. And, indeed, this theoretical prediction has been confirmed via extensive experimentation, most particularly through experimental tests of Bell's inequality.[2]

But this simple fact — that any two states of a composite system can be superposed — has utterly profound consequences. For example, since Anke and Bert can be in state $|\mathit{left}>_A|\mathit{left}>_B$ or in state $|\mathit{right}>_A|\mathit{right}>_B$, they can also be in the superposition state $|\mathit{left}>_A|\mathit{left}>_B + |\mathit{right}>_A|\mathit{right}>_B$. Let us call this superposition state $|E>$ for short. Then state $|E>$ says that if we perform a joint position-position measurement, we will always get the same result for both electrons; i.e. Anke and Bert always go through the same door.

But when the state of Anke and Bert is $|E>$, then what is Anke's state? Obviously, Anke is not in the state $|\mathit{left}>$, because $|E>$ says that it is possible for Anke to go through the right door. Similarly, Anke is not in the state $|\mathit{right}>$. So is Anke in the superposition $|\mathit{left}>+|\mathit{right}>$? No, that's not possible, because if Anke were in that state, then whatever state Bert is in, it would then be possible both that $|\mathit{left}>_A|\mathit{state}>_B$ and that $|\mathit{right}>_A|\mathit{state}>_B$. However, by replacing $|\mathit{state}>$ with either $|\mathit{left}>$, $|\mathit{right}>$, or a superposition thereof, you always get too many possibilities — you always get a state in which there is a chance of Anke and Bert going through opposite doors, which is inconsistent with $|E>$. Thus, the state $|E>$ rules out every possible one of Anke's states; when the composite system is in state $|E>$, then Anke is not in any state! All of Anke's quantities — position, velocity, etc. — lack determinate values. The pioneers of quantum mechanics invented a special name — "entanglement" — for situations like this where two objects are so intertwined with each other that they cease to have *any* individual characteristics.

So, quantum mechanics applies to small composite systems, such as a pair of electrons; and it predicts that such systems will have entangled states. In fact, there is nothing special about the number two; if we put together three or four electrons, we still get a system that obeys the laws of quantum mechanics. We might suppose, however, that this process of composition cannot go on indefinitely. At some point, we must reach a limit where quantum mechanics ceases to be valid. However, that supposition is false: despite many experiments, physicists have never found a cut-off point at which quantum mechanics ceases to be valid. In other words, all the evidence indicates that the composite of any two quantum-mechanical systems is another quantum mechanical system. Consequently, the laws of quantum mechanics hold for any objects, no matter how large or heavy, that are built out of other objects obeying the laws of quantum mechanics.

As a variation on a classic example, consider a cat called Tibbles. We can, in thought, build Tibbles up piece by piece from elementary particles. Beginning with two elementary particles A and B, which obey the superposition principle, we form a composite particle $A + B$, which then also obeys the superposition principle. We then add a third elementary particle, C, which of course obeys the superposition principle, and the result is a larger object $A + B + C$ that also obeys the superposition principle. Proceeding in this manner, we finally end up with Tibbles, a composite $A + B + C$ of elementary particles, who also is subject to the superposition principle. In particular, for any two states that Tibbles can be in, he can also be in the superposition of those two states.

Consider, for example, the state $|alive>$, in which Tibbles is alive, and the state, $|dead>$, in which Tibbles is dead. Then Tibbles can also be in the superposition state $|alive>+|dead>$, in which he is neither definitely alive nor definitely dead. Similarly, consider the state

$$|alive>_A|alive>_B+|dead>_A|dead>_B,$$

in which Tibbles is entangled with a mouse. Then Tibbles has no state at all, is neither alive nor dead, is not awake or asleep, etc. These consequences of quantum mechanics are not mere curiosities; they have utterly profound consequences for our understanding of physical reality.

Dynamics

The fact that cats can be in indeterminate states is hard to swallow. Believe it or not, though, there is an even more troubling consequence of quantum mechanics — namely it seems to show that when we "make observations" then we become entangled with physical objects, and so *we* end up having no determinate properties. This most troubling consequence of quantum mechanics is the result of combining the previous two postulates (superposition and entanglement) with the following simple fact about how quantum states change over time.

The theories of classical physics postulate the existence of deterministic dynamical laws. These laws provide a collection of conditional statements: if the state of the system at some earlier time is *S*, then the state of the system at some later time will be *S*'. Now, the situation in quantum mechanics is, in fact, the same: quantum states change in time according to the Schrödinger equation, which is completely deterministic in the sense that the current quantum state of an object determines uniquely its future quantum state. In addition, however, changes of quantum state always preserve superpositions. If the state |*S*> were to evolve into the state |*S*'>, and if the state |*T*> were to evolve into the state |*T*'>, then the state |*S*>+|*T*> would evolve into |*S*'>+|*T*'>. The assumption that superpositions are preserved through time is called "linear dynamics" or "unitarity."[3]

The most profound puzzle of quantum mechanics — namely, the measurement problem — is a result of linear dynamics in combination with the facts described in previous sections. Before we present the measurement problem, let's briefly summarize the features of quantum mechanics from which it follows.

- Superposition principle: any two possible states can be superposed. In a superposition state |*left*>+|*right*>, an object is neither in the state |*left*> nor in the state |*right*>; rather, its location is indeterminate.
- Entanglement: a pair of objects can be in an entangled state in which neither of the objects has any determinate properties.
- Linear dynamics: superpositions are maintained through dynamical changes.
- Size does not matter: the previous postulates apply to all physical objects, regardless of their size.

The measurement problem

Perhaps you can deal with the fact that electrons are in superposition states. Perhaps you can even accept that cats can be in entangled states in which they cease to have any properties whatsoever. After all, these are predictions of quantum mechanics, and we believe that quantum mechanics is true.

But why do we believe that quantum mechanics is true? We think it is true because it makes predictions, and these predictions are almost always correct. But how do we check these predictions? We check these predictions by making measurements — e.g. if quantum mechanics says that an electron has a 50 per cent chance of going through the left door, then we set up a detector to see how often it goes through the left door.

In order to measure whether the electron goes through the left door, we need some sort of detector. Let's suppose that there is a computer with sensors attached to both doors. If it detects an electron going through the left door, then it displays "left" on its monitor, and if it detects an electron going through the right door, then it displays "right" on its monitor. Suppose that before the computer detects anything, it displays "ready" on its monitor. Of course, the computer itself is a physical object, composed of stuff that obeys the laws of quantum mechanics. Hence, the computer must obey the laws of quantum mechanics — in particular, it must have superposition states, and it must be able to enter into entangled states with other physical objects.

Now we have just stipulated that the computer is a reliable detector of the door through which the electron travels. In other words, this means that if the initial state of the electron and computer is |*left*>|"*ready*">, then its final state will be |*left*>|"*left*">. Similarly, the initial state |*right*>|"*ready*"> leads to final state |*right*>|"*right*">.

But now let's check the prediction that quantum mechanics makes for when an electron is in the state |*left*>+|*right*>. (Recall that quantum mechanics predicts that in 50 per cent of the cases, the electron goes through the door on the left, and in 50 per cent of the cases it goes through the door on the right.) We begin then with state

(|*left*>+|*right*>)|"*ready*">,

which (since superpositions distribute over composition) is the same as the state

|*left*>|"*ready*">+|*right*>|"*ready*">.

We then ask, what will the state of the detector be after it interacts with the electron? But quantum mechanics (linear dynamics) tells us that superpositions are preserved through dynamical change, thus the final state of the electron and computer will be

|*left*>|*“left”*>+|*right*>|*“right”*>.

But this is an entangled state! In this state, the computer neither displays “left” nor “right”; in fact, the computer has become so entangled with the electron that it has no properties of its own. Our attempt to check the prediction has resulted in utter failure, since the computer displays nothing at all.

We have a mental problem

At this stage, you might be tempted to say: quantum mechanics makes a false prediction when it says that the computer ends up in an entangled state. In fact, we know (from experience) that the computer ends up either in the state |*“left”*> or in the state |*“right”*>.

But that is too fast. We do not know from experience that the computer ends up either in the state |*“left”*> or |*“right”*>. What we know from experience is that if we look at the computer monitor, then we will see either the state |*“left”*> or the state |*“right”*>. The fact that the computer ends up in an entangled state is consistent with our experience; in fact, it accurately predicts what our experience will be. Recall that a superposition state predicts equal probabilities for each of its components. But then if the computer and electron are in the superposition/entangled state

|*left*>|*“left”*>+|*right*>|*“right”*>,

we should expect that 50 per cent of the time |*left*>|*“left”*> will obtain, and 50 per cent of the time |*right*>|*“right”*> will obtain. But that prediction is accurate; that is what we do see when we perform this experiment.

If, however, we attempt to describe the observer herself in the language of quantum mechanics, then we face a serious problem. This problem is presented with force in the classical work *Quantum Mechanics and Experience*, by the philosopher David Albert. I quote his exposition at length (I use “left” and “right” in the place of Albert’s “hard” and “soft,” and I use a computer monitor in the place of a pointer):

> Suppose, then (just as we did before), that literally every physical system in the world (and this now includes human beings; and it includes the brains of human beings) always evolves in accordance with the dynamical equations of motion . . . Being a “competent observer” is something like being a measuring device that’s set up right: What it means for Martha to be a competent observer of the computer monitor is that whenever Martha looks at a monitor that displays “left”, she eventually comes to *believe* that the monitor displays “left”; and that whenever Martha looks at a monitor that displays “right”, she eventually comes to believe that the monitor displays “right”. What it means (to put it more precisely) is that the dynamical equations of motion entail that Martha (who is a physical system, subject to the physical laws) behaves like this:
>
> $$|\mathit{ready}>_{o}|\mathit{ready}>_{m} \rightarrow |\text{“}\mathit{ready}\text{”}>_{o}|\mathit{ready}>_{m},$$
> $$|\mathit{ready}>_{o}|\text{“}\mathit{left}\text{”}>_{m} \rightarrow |\text{“}\mathit{left}\text{”}>_{o}|\text{“}\mathit{left}\text{”}>_{m},$$
> $$|\mathit{ready}>_{o}|\text{“}\mathit{right}\text{”}>_{m} \rightarrow |\text{“}\mathit{right}\text{”}>_{o}|\text{“}\mathit{right}\text{”}>_{m}.$$
>
> In these expressions, $|\mathit{ready}>_{o}$ is that physical state of Martha’s brain in which she is alert and in which she is intent on looking at the monitor and finding out what it says; $|\text{“}\mathit{ready}\text{”}>_{o}$ is that physical state of Martha’s brain in which she believes that the monitor displays the word “ready”, [etc.] . . .
>
> Let’s get back to the story. The state of the electron and the computer is the strange one $|\text{“}\mathit{left}\text{”}>_{m}|\mathit{left}>_{e}+|\text{“}\mathit{right}\text{”}>_{m}|\mathit{right}>_{e}$. And now in comes Martha, and Martha is a competent observer of the computer monitor, and Martha is in her ready state, and Martha looks at the monitor. It follows from the linearity of the dynamical equations of motion (if those equations are right), and from what it means to be a competent observer of the monitor, that the state when Martha’s done is with certainty going to be
>
> $$|\text{“}\mathit{left}\text{”}>_{o}|\text{“}\mathit{left}\text{”}>_{m}|\mathit{left}>_{e}+|\text{“}\mathit{right}\text{”}>_{o}|\text{“}\mathit{right}\text{”}>_{m}|\mathit{right}>_{e}.$$
>
> That’s what the dynamics entails.
>
> . . . That state described [in the prior paragraph] is at odds with what we know of ourselves by *direct introspection*. It’s a superposition of one state in which Martha thinks that the monitor displays “left” and another state in which Martha thinks that the monitor displays “right”; *it’s a state in which there is no matter of fact about whether or not Martha thinks the monitor displays anything in particular*.
>
> And so things are turning out badly.[4]

Thus, according to Albert, quantum mechanics entails a fact — that at the end of a measurement, a person will not have any belief about the outcome — that would utterly destroy our ability to test the predictions of quantum mechanics. Therefore, quantum mechanics is incoherent.

Interpretations of quantum mechanics

Superposition states and entangled states might be puzzling, and they certainly require a stretch of our conceptual framework. But the *measurement problem* is not merely a puzzle; rather, it is an apparent proof of the incoherence of quantum mechanics. We must do something to save quantum mechanics from incoherence; otherwise, the best physical theory in history is shown to be a sham, and certainly not worth your attention as a guide to understanding the nature of reality.

An *interpretation* of quantum mechanics is an attempt to explain how it could be true. But of course, if the theory leads to a contradiction then it cannot possibly be true. Thus, an interpretation of quantum mechanics must reject or modify one of the assumptions that was used to derive the measurement problem.

Since the origin of quantum mechanics, there have been a number of responses to the measurement problem, and these responses can be classified according to which assumption they reject. First, some interpretations of quantum mechanics (so-called "hidden variable interpretations") solve the measurement problem by rejecting the superposition principle. Second, some interpretations of quantum mechanics (especially dynamical reduction theories) solve the measurement problem by rejecting the dynamical laws of quantum mechanics. Third, some interpretations of quantum mechanics (especially Everettian or *many worlds* interpretations) solve the measurement problem by denying that observation really does occur in the sense we normally suppose it does. In the remainder of this section, I will give a brief overview of each of these interpretive strategies. In the following section, we will explore the interpretation of quantum mechanics in light of a conscious presupposition that human beings are more than just hunks of physical matter.

First, the measurement problem can be blocked by denying that there are superposition states in which some quantities are indeterminate. Nobody denies that the state |*left*>+|*right*> is possible. However, we are not *necessitated* into saying that it is a state in which the electron has no position. The argument that the electron lacks a position is roughly as follows: if the electron had some position (but we didn't know which), then we would not get a diffraction pattern on the screen; rather, we would get the two lumps behind the doors. But that argument is not air-tight: the conclusion only follows if we assume that the electron is a classical particle not subject to any additional forces or laws. It remains a possibility that a particle-like entity could produce a diffraction pattern if it obeyed laws of motion that were quite different from the laws discovered by Newton.

The strategy outlined above is sometimes misleadingly called giving a hidden variable interpretation of quantum mechanics, the most famous example of which is the theory developed by David Bohm, now called Bohmian mechanics.[5] While this strategy promises to solve the measurement problem while maintaining determinism (which some find desirable), it also has several difficulties. Most notably, Bohmian mechanics postulates the existence of a guiding field of somewhat questionable metaphysical credentials. In particular, the guiding field (unlike all the other physical fields we know and love) carries no energy-momentum, and so is empirically undetectable. Furthermore, the components of the guiding field are not associated with localized regions of space in the way that, say, the electromagnetic field is. Thus, the guiding field is not a field in the traditional sense, and in particular it does not play the traditional role of a field as a mediator of cause and effect relations in space and time. The mysterious nature of the guiding field was itself an insuperable obstacle for Einstein (who otherwise longed to replace quantum mechanics with a deterministic theory). On the other hand, Bohm himself proposes a new metaphysics in which the guiding field is itself a quasi-mental entity.

> These new properties suggest that the field may be regarded as containing objective and active information, and that the activity of this information is similar in certain key ways to the activity of information in our ordinary subjective experience. The analogy between mind and matter is thus fairly close. This analogy leads to the proposal of the general outlines of a new theory of mind, matter, and their relationship, in which the basic notion is participation rather than interaction.[6]

Bohm's ideas might sound intriguing, but they are far from metaphysically innocent. We might wonder: does quantum mechanics require a radically new theory of mind and matter, or is it consistent at least in general outlines with the wisdom handed down to us through the ages?

Second, some physicists blame the measurement problem on the dynamical laws of quantum mechanics, and in particular on quantum mechanics' supposition that superpositions are preserved over time.

According to these physicists, quantum mechanics is simply a *false* theory, and needs to be replaced by a different theory. Moreover, these physicists have gone on to provide concrete proposals for these alternative theories. The most famous alternative to quantum mechanics is the dynamical reduction theory proposed by Ghirardi, Rimini, and Weber.[7] According to the GRW theory, quantum mechanics is *usually* right about how things change over time. However, once in a blue moon, there is a random and spontaneous collapse of the state of an object. For example, there is an extremely small probability that an electron in state |*left*>+|*right*> will spontaneously transition into state |*left*> or |*right*>. This probability is so small that it is extremely unlikely that an individual electron's state would collapse, even over the entire history of the universe. However, in order to solve the measurement problem, GRW take advantage of the fact that collapses are contagious; i.e. if one particle is entangled with another, and if the state of the first collapses, then the state of the composite automatically collapses. But it follows, then, that for a system consisting of a very large number of particles — e.g. a measuring device — there is a non-negligible probability that one of its constituent particles will collapse, and hence that the state of the big composite object will collapse. So, GRW dynamics would explain why quantum mechanics is *almost* true for very small objects, but often false (since wave functions collapse) for large objects.

Finally, some propose to solve the measurement problem by rejecting the intuition that a measurement ends with one definite outcome to the exclusion of the other possibilities — in particular, by rejecting the claim that a reliable observer will believe either that the computer monitor shows "left" or "right." The most famous version of this strategy — alternately called the Everett interpretation or the many worlds interpretation — was introduced by Hugh Everett.[8] According to Everett, when a person makes an observation or measurement she becomes entangled with the measuring device and with the object under study. Thus, at the end of the measurement, the person is not in the state "I believe that the monitor displays 'left'" and she is also not in the state "I believe that the monitor displays 'right'."

But why then do we mistakenly believe that we often have definite perceptual beliefs? Proposed answers to this question are, by necessity, sophisticated and involve serious grappling with the metaphysics of the mind-body relation. (To follow some recent developments, one might look at the work of the physicist Don Page of the University of Alberta, or of the philosophers Hilary Greaves, Simon Saunders and David Wallace of Oxford University.) In short, the Everett interpretation proposes that when a person makes a measurement, then the universe itself splits into several branches, and the person making the measurement is split into several copies of herself.

Now, the idea of a branching universe is not, in and of itself, so metaphysically absurd that it counts decisively against the Everett interpretation. Indeed, a branching universe would be a natural way to cash out the idea that the future is open and not determined by the past. If that was the only metaphysical revision required, then I might be tempted to recommend the Everett interpretation to you. In fact, there are versions of the Everett interpretation that are explicitly consistent with mind-body dualism, namely the single mind and many minds interpretations of David Albert and Barry Loewer.[9]

But, unfortunately, the naive *many worlds* version of Everett's interpretation, as well as the single and many minds interpretations, are vulnerable to a number of objections that have been cataloged over the past thirty years. For example, in the many worlds interpretation, the universe is supposed to split into many parts when a measurement occurs. But how can a measurement in one place cause a change in the entire universe, including very distant regions, without violating the laws of special relativity? Furthermore, if all the possible measurement outcomes are actualized (in some universe, or relative to some mind), then what sense does it make to say that certain outcomes are more likely than others?[10]

I will not claim that these problems with the Everett interpretation cannot be solved. Indeed, an extremely clever and philosophically cogent version of the Everett interpretation has been developed recently by the Oxford University philosophy of physics group. But this recent work makes it clear that the Everett interpretation is no friend of mind-body dualism. Indeed, the Everett interpretation is most plausibly and compellingly developed in the context of a form of "functionalist physicalism."[11] Thus, while a physicalist may have good reasons to look to the Everett interpretation as a key to understanding physical reality, a dualist has just as good a reason to look elsewhere.

Each of these interpretations of quantum mechanics agrees that there is a problem that needs to be solved. The first two interpretations solve the problem by rejecting an assumption of quantum mechanics — in one case the assumption that superpositions entail indeterminacy, and in the other case the assumption that superpositions are preserved through time. The third interpretation solves the measurement problem by revising our intuitive idea about what happens when we make observations or measurements. Each strategy has its virtues and its drawbacks. However, none of the strategies takes seriously the idea that "observation" involves a non-physical thing (e.g. a mind or a soul). Some might claim that it is a virtue of these interpretations that they need not assume the existence of non-physical things. But if you already and independently believe in the existence of non-physical things, then there is no good reason to forget this fact when interpreting quantum mechanics.

Quantum mechanics on the Soul Hypothesis

I claim that a dualist should be wary of the textbook derivation of the measurement problem (as, for example, in Albert's book), because this derivation relies on a tacit assumption of reductionist physicalism. In particular, Albert tacitly assumes physicalism when he says that "|"*left*"> is that physical state of Martha's brain in which she believes that the computer monitor displays the word 'left'." (Note how the first part of the sentence is about a physical feature of Martha's brain and the second part of the sentence is about Martha's mental state.) According to dualism, there are two states in play here: there is the state of Martha's brain, and there is her mental state. So, Albert is using one name for what the dualist claims are two different things; in other words, he is tacitly equating mental states with physical states. But if we distinguish the two sorts of states, then it is not obvious that the argument for the measurement problem will go through.

In the remainder of this chapter, I reexamine the measurement problem in light of the fact that human observers are not just chunks of physical matter. First, I argue that mental states, unlike physical states, cannot be superposed, and therefore cannot become entangled with physical states. This point itself would be sufficient to block the derivation of one half of the measurement paradox — the claim that an observer fails either to see "left" or to see "right." But I will go further; in order to demonstrate, beyond a doubt, the coherence of quantum mechanics and the Soul Hypothesis, I suggest a model of the interaction of physical states and mental states relative to which mental states reliably track states of the physical world.

The two state space hypothesis

The Soul Hypothesis is, of course, a pre-theoretical idea in the sense that the statement "human beings are more than just their bodies" is not yet precise enough to bring to bear directly on the question of how we should describe a person's mental states when she is performing measurements on objects like electrons. So, if we are to say something concrete about the interaction of physical and mental states, then we must — with all due humility! — try to translate the Soul Hypothesis into something like a precise metaphysical thesis.

Many philosophers throughout history have proposed and defended precise versions of the Soul Hypothesis. I will not, in this chapter, try to survey the various proposals, or elaborate on my choice of a proposal. Suffice to say that the Logical Independence Hypothesis, a precisification of the Soul Hypothesis, has been defended by several notable dualists.[12] It holds that mental states are *logically* independent from physical states. That is, for any possible mental state, and any possible physical state, it is possible that the two states could obtain at the same time.

Of course, there are stable correlations in our world between physical states and mental states, and so there are probably laws of nature connecting the two. But the point of the logical independence hypothesis is that the two sorts of states are *conceptually* distinct — a physical state is a different sort of thing than a mental state, and physical states do not *logically* or *conceptually* necessitate mental states, and vice versa.

The sort of independence that the dualist postulates between mental states and physical states is just like the mix and match principle that holds between distinct physical quantities (e.g. position and velocity in classical physics), or between quantities of distinct physical objects (e.g. the position of Jupiter and the position of Mars). Thus, if $|M>$ is a mental state, and $|P>$ is a physical state, then we can borrow from quantum mechanics the notation $|M>|P>$ to denote the conjunctive state whose possibility is asserted by the independence thesis. But this notational adjustment is far from trivial: if we have distinct names for physical states and mental states, then obviously conclusions about physical states (e.g. they can be superposed) cannot be *automatically* transferred to mental states.

The non-superposability of mental states

Quantum mechanics entails that Martha's brain can be in a superposition of states. But if Martha's mental states are not identical to her brain states, then it does not immediately follow that she can be in a superposition of mental states. In fact, I claim that mental states cannot be superposed. I will back this claim up both by pointing out a lack of evidence for their superposability, and by providing positive arguments against the superposability of mental states.[13]

Why do we think that physical states can be superposed? The answer is *not* that we *see* that one state is a superposition of two other states

— indeed, we have no idea what that would look like. Rather, superposition is an unobservable, theoretical relation between states; and this relation was postulated because it explains phenomena (e.g. the two-slit experiment). The postulation of unobservable structure, behind the phenomena, is a common strategy of theoretical science; its justification comes from the fact that it explains empirical facts that would otherwise be puzzling. For example, Einstein's theory of relativity postulates that space and time have hidden geometrical structure; this postulate is justified by the fact that it explains the motions of planets and stars. But are there phenomena for which we would have an explanation if we posited the existence of an unobservable relation of superposition between mental states?

If we can trust the scientific experts (namely, psychologists), then the answer is no: psychologists have not postulated the existence of superpositions of mental states, and indeed they have found no use for this concept. But we can make the argument even stronger. What makes the concept of superposition as found in quantum mechanics scientifically acceptable is the fact that quantum mechanics provides the means to identify which states are superpositions of which other states (e.g. superpositions of spin-*x* states are spin-*y* states); and moreover it describes the empirical manifestations of superposition states (e.g. the superposition of |*left*> and |*right*> manifests a diffraction pattern). In other words, superposition is not an empty concept, but a concept with testable empirical content. But now let's apply this sort of rigorous standard to mental states. Consider the state of your mind when you see "left" on the computer monitor, and the state of your mind when you see "right" on the computer monitor. Now, if someone claims that these two states can be superposed, then he should be able to back this claim up by identifying the resulting state, and describing that state's empirical manifestations. Otherwise, his claim that such a state exists is empty, and does no explanatory work. But nobody has the first clue how to identify superpositions of mental states; indeed, no serious scientist has even ventured a speculative theory of the superposition of mental states. So, the claim that there are superpositions of mental states cannot be taken to be a serious scientific claim.

What, in contrast, might somebody say to argue for the existence of superpositions of mental states? The only possible argument I can imagine would be an argument by analogy: all physical states can be superposed, therefore (in absence of further evidence) we should suppose that mental states can be superposed. But why should we think that what's true of *physical* states should also be true of *mental* states

— unless of course, we have already decided that mental states are nothing but physical states in disguise? Perhaps the defender of mental superpositions will claim that if mental states are to be correlated with physical states, then these mental states will themselves need to be superposable. But that supposition is provably wrong: in what follows, I will show that physical states and mental states can be strictly correlated, even though the latter cannot be superposed.

The non-existence of mental superpositions has profound consequences for the states of a composite mental-physical system. In particular, according to quantum mechanics, composite physical objects can enter into entangled states in which neither individual object has any determinate properties — and this is precisely what was shown to happen in a measurement. But entanglement requires superposition: if states cannot enter into superpositions, then they also cannot become entangled.

INTERACTIONIST DYNAMICS

To this point I have argued — based on the assumption that mental states are distinct from physical states — that mental states are not superposable, and that mental states cannot become entangled with physical states. These points are enough to block the derivation of the measurement problem: they block the derivation of the claim that Martha fails, at the end of a measurement, to be in a state of seeing that something is so.

To defend the coherence of quantum mechanics against the measurement problem (in particular, to show that it does not entail a contradiction), it is fully sufficient to uncover an error (or tacit, but false, assumption) in the proof of one of the contradictory claims. But Albert's derivation of the measurement problem tacitly assumes physicalism, in contradiction with the starting point of this chapter and of this book as a whole. So, we would be fully justified in concluding this chapter at this point, having noted that the most severe problem for quantum mechanics emerges from an overly simplistic view of the mind-body relation.

But we always want to know more, in particular how mental and physical states might interact in such a way that we (conscious observers) are able reliably to gain information about the external world. Accordingly, I will proceed to sketch some ideas that might lead to a coherent understanding of how mental and physical states interact when we make observations, and in particular observations of quantum mechanical objects.

Consider again the computer monitor with its two states |*left*> and |*right*>. Let |"*left*"> be Martha's mental state in which she believes that the computer screen displays "left," and let |"*right*"> be Martha's mental state in which she believes that the computer screen displays "right." If Martha is a reliable observer, then an initial state |*ready*>|*left*> should lead to the final state |"*left*">|*left*>, and an initial state |*ready*>|*right*> should lead to the final state |"*right*">|*right*>. But now suppose that the initial state is:

$$|ready\rangle(|left\rangle+|right\rangle) = |ready\rangle|left\rangle+|ready\rangle|right\rangle.$$

If Martha's mental states could become entangled, then we would expect the final state to be an entangled state — that would follow from the assumption of linear dynamics (i.e. that superpositions are preserved through changes). But the resulting entangled state is *not possible*. We cannot apply the requirement of linear dynamics if it would lead to an impossible state.

In fact, it is impossible to fill out the story of what happens to Martha and the computer using deterministic dynamical laws. That is, if the computer starts out in state |*left*>+|*right*>, then the future state of Martha and computer is not determined: sometimes it will be |"*left*">|*left*>, and sometimes it will be |"*right*">|*right*>. Indeed, if we were to measure the initial state of Martha and the computer, then in 50 per cent of cases it would yield |*ready*>|*left*>, and in 50 per cent of cases it would yield |*ready*>|*right*>. Furthermore, we stipulated that |*ready*>|*right*> would lead to |"*right*">|*right*>, and similarly |*ready*>|*left*> would lead to |"*left*">|*left*>. Thus, applying the principle (as in common sense and classical physics) that probability is preserved through time, the final state should predict |"*right*">|*right*> in half of the cases, and |"*left*">|*left*> in the other half of the cases. However, because Martha's mental states cannot become entangled with the computer, there is *no* state that makes this prediction. Therefore, the future state of Martha and the computer cannot be determined by its initial state.

Could there then be indeterministic, or probabilistic dynamical laws that govern both aspects of the universe — physical and mental — and their interaction? From a purely mathematical point of view, there certainly could be. Indeed, the difficulty at this point is that we have *too many* options, and not enough evidence to choose between them.

First, we already have a dualist-friendly interpretation of quantum mechanics in the work of Henry Stapp.[14] But since Stapp has already written extensively and accessibly on his approach to quantum mechanics, I leave it to the diligent reader to explore these ideas on his or her own. Thus, I conclude this chapter by mentioning a few more of the possible options a dualist has for interpreting quantum theory.

First, one could take the equations of motion of Bohmian mechanics and reinterpret the terms referring to determinate particle configurations as referring to determinate mental states, in which objects are observed to be in determinate locations in space. (The resulting theory might be similar to Albert and Loewer's single mind theory.[15]) However, the resulting theory might not be the most natural for traditional interactionist dualism, because the theory would seem to endow perceptual states with their own autonomous dynamics rather than making them responsive to the states of the external world.

Second, and more promisingly, the Ghirardi-Rimini-Weber (GRW) collapse theory solves the measurement problem by introducing indeterministic dynamical laws. But one problem with GRW is that it seems to lack independent motivation: the collapse dynamics seems to be ad hoc, and put in by hand to solve the measurement problem. But here the dualist may have an advantage. In particular, we live in a universe with two types of things (physical and mental) with different natures; in particular, the physical things have superposition states, but the mental things do not. Now, suppose that the "natural" dynamics of physical things are the laws of quantum mechanics. However, we have seen that if a physical thing (e.g. a brain) is joined to a non-physical thing (e.g. a mind) in such a way that their states are correlated in a law-like way, then the physical thing *cannot* exactly and without exception obey the laws of quantum mechanics. (The non-existence of superpositions of mental states entails that the joint physical-mental object cannot obey the laws of quantum mechanics.) But what then is the next best thing? If the physical part in isolation would follow the rules of quantum mechanics but is constrained by the nature of its mental counterpart, then the GRW laws would provide a highly natural and harmonious way for these two sorts of objects to interact with each other and with other physical objects. Thus, a dualist could happily follow (or contribute to) the development of the GRW theory, but could underwrite it with independent motivation coming from his or her background metaphysical framework.

In conclusion, the sciences rightly take a central place in our efforts to develop an accurate system of beliefs. After all, the sciences are nothing more than a systematic effort to submit our beliefs to the tribunal of the external world. But the example of the measurement problem shows poignantly that it is naive or disingenuous to claim to approach the data from a standpoint of metaphysical neutrality, and to expect the data to provide its own interpretation. Rather, we always see the world through the lens of our background metaphysical assumptions; and if we put bad metaphysics into our scientific theories, then we can expect to get bad metaphysics out of them. (And, tragically, people of common sense sometimes throw out the baby with the bath water: they blame the scientific theories themselves rather than the interpretative supplements to these theories.) In the case of quantum mechanics, if one presupposes physicalism, then one quickly lands in the measurement problem; and one may then say crazy things about a new metaphysics of unfolding conscious wavefunctions, or minds being nothing but functional patterns in a universal wavefunction, or there being no objective reality outside of our perceptions. In contrast, if one begins with a common sense assumption of dualism, then one finds no reason in quantum mechanics to reject this assumption; quite to the contrary, quantum mechanics proves to be surprisingly in harmony with the accumulated wisdom of our metaphysical and scientific forebears.[16]

As we move on to our next essay, we point out that the two-slit experiments that Halvorson describes show that people's observations — the perceptions made by their souls — influence the outcomes of experiments in the laboratory. If that is true, is it really any more incredible that the decisions of their souls influence the outcomes of neurons firing in their brains, as assumed for example in Goetz's essay? We now know experimentally that more incredible things than this are happening all around us.

Halvorson has led us back from the physical to the psychological by way of the measurement problem — the role that people's observations play in current physics. With Dean Zimmerman's essay, we return entirely to the psychological level of our more everyday experience, considering again in more detail the implications of the (we believe) undeniable fact that people have perceptual experiences.

Like Halvorson, Zimmerman adopts the assumption that a psychological property of me like "I am seeing a patch of red over there" is (at least) logically distinct from any physical property that I have, such as my various brain states. Indeed, Zimmerman fills in briefly some of the reasons why many contemporary philosophers make this assumption, which Halvorson simply assumes. For example, it seems logically possible that there could be creatures that behave just like we do, but have no inner life, no first-person experience of qualia at all — what philosophers refer to as "zombies." Or it seems possible that there could be creatures just like us who have perceptual experiences, but whose experiences are systematically different from ours: for example, maybe when they see something that gives me the quale that I call red, they actually have a quale that I would call green, whereas what I experience as green they experience as red. If such differences are possible, then it seems that psychological states are partially independent of physical states. This leads to a view that is known among the professionals as "property dualism"—the idea that there are fundamentally different kinds of properties that something like a person might have. Much as a single object like a ball can have two different physical properties at once (say, being red and being hard), so I might have two different kinds of properties at once: the physical property *of having a certain kind of cone fire on the retina of my eye, and the psychological property of experiencing red.*

But instead of considering the implications of the distinction between physical and mental properties for quantum physics, Zimmerman considers its implications for the Soul Hypothesis more directly. Many people in the current intellectual scene are open to the idea of property dualism (that there are two distinct kinds of properties, physical properties and psychological ones) who are not open to traditional substance dualism — the idea that there are two distinct kinds of things, physical things like bodies and non-physical things like souls. Mere property dualism no doubt seems like a safer, more conservative hypothesis to them. But Zimmerman presents a philosophical line of argument that it is harder to take this easy way out than it might appear at first.

Zimmerman's argument centers on the issue of vagueness. This is the undeniable fact that all middle-sized objects that we are familiar with have imprecise boundaries in both time and space. It is not perfectly clear just where they start or stop, or exactly when they first begin to exist or stop existing. As I hiked from the level valley up to the peak of the mountain, at what point did I first set foot on the mountain itself? As an acorn in my yard germinated and sprouted and grew large and strong, at what instant did it first become a tree? We do not expect precise, non-arbitrary answers to questions like this. But our simplest and most fundamental experiences are not vague and fuzzy in the same way that these physical objects are. How then can a vague and imprecise physical object, like my brain or my nervous system or my whole body, directly produce the discrete and unique experiences that I have? Zimmerman suggests that careful thinking about this should free us from the notion that property dualism is really simpler and more conservative than substance dualism. On the contrary, he suggests that it is much more plausible that new kinds of properties like "seeing red" exist because new kinds of things exist that have those properties — namely souls.

In essence, Zimmerman's argument is as follows. The vague objects that are the standard candidates for what a person fundamentally is according to materialism are a brain, or a nervous system, or a human body, or the like. These are the obvious candidates because (i) they can be identified as distinguishable units by (say) an anatomist, and (ii) the direct causes of mental events seem to be located inside them. But it turns out that vague objects like brains and bodies are not suitable candidates for having mental properties because the laws of nature that presumably link physical properties and mental properties in systematic ways must be precise. After all, the fundamental physical laws that we know of refer *to precise things like electrons and photons, not to imprecise composite things like mountains and trees, and crucially so. It follows, then, that the subject of basic mental properties must also be precise in nature, and human brains and bodies do not fit the bill. (Or, even if the mental properties are not entirely precise, there is no reason to think that any vagueness they might have corresponds directly to the familiar kind of vagueness that bodies and brains have.) One prevalent way to think about vague objects like mountains and trees is to say that there is a veritable host of objects in the vicinity that could be the precise mountain or the precise tree that we refer to in a particular situation. If we said the same thing about human brains and bodies, then there would also be a veritable host of objects (e.g. slightly different collections of cells or molecules or atoms) in the vicinity that could be the precise brain or body that is the subject of mental properties. But none of these arbitrarily chosen but precise objects presents itself as the better candidate for being the precise subject of mental properties. In the face of this quandary, then, Zimmerman bids us remember that substance dualism provides a natural alternative: that we are in essence souls, and not some imprecise physical thing. Since souls are (by hypothesis) not physical things, they are not a composite of cells, molecules, and atoms. Therefore, their boundaries in space and time need not be vague in the ways that medium scale physical objects necessarily are. They could then be precisely the right things to have the psychological properties that we know we have.*